(2)

Office of Naval Research contract no. N00014-88-K-0377

The Concept of Best Probability in the Analysis of Approximation Algorithms

Dorit S. Hochbaum, Principal Investigator

FINAL TECHNICAL REPORT

This is to report on the research findings supported by ONR grant N00014-88-K-0377. The research is described in 19 research papers. The presentation here summarizes the findings in each paper. The period covered is July 1988 through May 1991.

I would like to take this opportunity to express my gratitude for providing me with this support. It has certainly made my work substantially more productive. In addition I wish to thank ONR for the support provided during this period to my consultants/co-authors, R. Shamir and S. Cosares. This was an essential factor in making the work possible.

The papers are listed in two categories, publications and working papers. The publications section includes all papers that either appeared in print or were accepted for publication. The section of working papers includes completed papers that are in status of revision or review.

PUBLICATIONS

1. "A Strongly Polynomial Algorithm for the Quadratic Transportation Problem with Fixed Number of Suppliers" (with S. Cosares), manuscript, August 1990, to appear in *Mathematics of Operations Research.*

The Transportation problem with a linear objective function is known to be solvable in strongly polynomial time, whereas instances with a convex non-quadratic objective function are not, even for cases with the integrality constraints relaxed. This paper presents a linear time algorithm for the Continuous Quadratic Transportation problem (QTP) with two source nodes. Further, it is shown how problem instances with a fixed number of source (or destination) nodes, k, could be solved in strongly polynomial time, $O(n^{k+1})$, using an algebraic tree computation model. The algorithms exploit a relationship between the Transportation problem and the Resource Allocation problem. The strong polynomiality of the algorithms presented here imply the existence of strongly polynomial algorithms for Integer Quadratic Transportation problems with a fixed number of source nodes as well.

2. "An Exact Sublinear Algorithm for the Max-Flow, Vertex Disjoint Paths and Communication Problems on Random Graphs," August 1988, to appear in *Operations Research.*

We describe a randomized algorithm for solving the maximum flow and minimum cut problems and the vertex disjoint paths problem on random graphs. The problems come up in identifying communication routes that do not overlap. The algorithm is very fast - it does not even look up most nodes in the network. Another feature of this algorithm is that it almost surely provides, along with an optimal solution, a proof of its optimality. In addition, the algorithm's solution is by construction a collection of vertex disjoint paths that is maximum. With some restriction on the graph density, an optimal solution to the NP-hard communication problem - finding a maximum collection of vertex-disjoint paths between sender-receiver pairs of terminals - is provided as well. The algorithm lends itself to a sublogarithmic parallel and distributed implementation. The effectiveness of the algorithm is demonstrated via extensive empirical study.

3. "Why Should Biconnected Components Be Identified First," to appear in, *Discrete Applied Mathematics.*

Most graph optimization problems are solved on each connected component of the graph separately. This requires the identification of the connected components of the graph. We show here that for several graph optimization problems, including the weighted vertex cover and the independent set problems, it suffices to know how to solve the problem on each **biconnected** component of the graph. The additional work required to give a solution on the whole graph takes a linear additive factor at most, whereas the potential savings in total running time are substantial. The same approach applies to approximation algorithms, and the approximation error bound is at most the maximum error bound among the biconnected components.

4. "The Use of Heuristics, Cutting Planes and Subgradient Optimization for the Multicovering Problem" (with N. Hall), to appear in, *European Journal on Operational Research.*

The multicovering problem is an extension of the set covering problem, where each element requires more than one unit of coverage. This is frequently the case where the covering requires certain redundancy in order to increase reliability. It is also the case when there are elements of identical features so that each requires separate coverage. The problem has many applications in scheduling, locations and business areas, and it arises in commercial or military contexts. The problem is even more intractable than the notoriously hard set covering problem. Here we describe a special purpose algorithm where various computational tools are put together and the structure of the problem is exploited.

The paper describes the algorithm and gives computational results for randomly generated problems.

5. "A Polynomial Algorithm for an Integer Quadratic Nonseparable Transportation Problem" (with J. G. Shanthikumar and R. Shamir), to appear in, *Mathematical Programming.*

In this paper we consider a quadratic transportation problem that arises as a high-multiplicity scheduling problem. The scheduling problem, minimizing the total weighted tardiness of jobs on a single machine , is NP-complete. The case analyzed in this paper is when all jobs are of unit length, there are numerous jobs, yet only a small number of job sets. In each job set, all jobs have identical due date and penalty weight parameters. Due to the large sizes of these sets, or their multiplicities, an algorithm which solves the unit-job problem with distinct jobs is only pseudo-polynomial and therefore prohibitively inefficient. We are thus seeking an algorithm whose number of operations will be polynomial, and, if possible, independent of the sizes of the sets.

It is shown that this scheduling problem can be reformulated as an integer quadratic nonseparable transportation problem. We provide a polynomial algorithm for the problem. When all weights are distinct, we also provide a polynomial algorithm which solves the problem in running time independent of the multiplicities and the due-dates. We employ methods of real analysis and prove a tight proximity result between the integer solution to the problem and a fractional solution to a related problem. A straightforward rounding algorithm is then used to derive the optimal integer solution from the fractional solution, in running time which is independent of the sizes of job sets.

6. "On the Impossibility of Strongly Polynomial Algorithms for the Allocation Problem and its Extensions," *Proceedings of the Integer Programming and Combinatorial Optimization Conference*, May 1990, 261- 274.

This paper describes results that complement the ones in papers 8 and 9. Whereas in those papers it is established that convex (concave) separable optimization problems over linear constraints are solvable in polynomial time (that depends on the magnitude of the largest subdeterminants), it is shown here that these polynomial algorithms are in a sense the best possible to achieve. This paper establishes a lower bound on the compexity of such problems as convex separable network optimization problem. It is shown that it is impossible to derive algorithms for such problems (as well as many other nonlinear problems) that run in strongly polynomial time, i.e. running time that is independent of the objective function coefficients or the right hand sides of the constraints. This is the first result in the literature demonstrating the impossibility of strongly polynomial algorithms.

7. "Strongly Polynomial Algorithms for the High Multiplicity Scheduling Problems" (with R. Shamir), *Operations Research*, Vol. 39, No. 4, 1991, 648-653.

A high multiplicity scheduling problem consists of many jobs which can be partitioned into relatively few groups, where all the jobs within each group are identical. Polynomial - and even strongly polynomial - algorithms for the standard scheduling problem (in which all jobs are assumed to be distinct) become exponential for the corresponding high multiplicity problem. In this paper we study various high multiplicity problems of scheduling unit-time jobs on a single machine. We provide strongly polynomial algorithms for constructing optimal schedules with respect to several measures of efficiency (completion time, lateness, tardiness, the number of tardy jobs and their weighted counterparts). The algorithms require a number of operations polynomial in the number of groups rather than in the total number of jobs.

As a by-product, we identify a new family of nxn Transportation problems which are solvable in $O(n \log n)$ time by a simple greedy algorithm.

8. "Convex Separable Optimization is Not Much Harder than Linear Optimization" (with J.G. Shanthikumar), *Journal of ACM*, Vol. 37, No. 4 (1990), 843-862.

In this paper we prove the polynomiality of nonlinear separable convex (concave) optimization problems, on linear constraints with a matrix with "small" subdeterminants, and the polynomiality of such integer problems provided the integer linear version of such problems is polynomial. We present a general purpose algorithm for converting procedures that solves linear programming problems with or without integer variables, to procedures for solving the respective nonlinear separable problems. The conversion is polynomial for constraint matrices with polynomially bounded subdeterminants. Among the important corollaries of the algorithm is the extension of the polynomial solvability of integer linear programming problems with totally unimodular constraint matrix, to integer separable convex programming problems. We also present an algorithm for finding an ε-accurate optimal continuous solution to the nonlinear problem which is polynomial in $\log\frac{1}{\varepsilon}$ and the input size and the largest subdeterminant of the constraint matrix. These developments are based on proximity results between the continuous and integral optimal solutions for problems with any nonlinear separable convex objective function. The practical feature of our algorithm is that it does not demand an explicit representation of the nonlinear function, only a polynomial number of function evaluations on a prespecified grid.

9. "The Complexity of Nonlinear Optimization" (with J. G. Shanthikumar), *Lecture Notes in Computer Science*, 372, Springer-Verlag, July 1989, pp. 461-472.

This paper describes classes of nonlinear optimization problems over linear constraints, and establishes their polynomiality. A complexity framework is introduced for the purpose of describing the output of such problems on digital computers. This is an early version of paper 8.

10. "Asymptotically Optimal Linear Algorithms for the Minimum K-Cut in a Random Graph" (with O. Goldschmidt), *SIAM J. on Discrete Mathematics*, February 1990, Vol. 3, No. 1, pp. 58-73.

The k-cut problem is to find a partition of a graph into k nonempty components, such that the number of edges between components is minimum. We prove that, when k is fixed, a k-cut of almost every random graphs consists of $k-1$ isolated vertices and one component on the remaining $n-k+1$ vertices. An important outcome of this property is a linear algorithm that derives the minimum k-cut in such graphs, almost certainly.

11. "A Fast Perfect Matching Algorithm in Random Graphs" (with O. Goldschmidt), *SIAM J. on Discrete Mathematics*, February 1990, Vol. 3, No. 1, pp. 48-57.

The matching problem is to find a maximum collection of mutually non adjacent edges in a graph. We present an algorithm that delivers a perfect matching in a random graph almost surely. The expected running time of this algorithm is $O(n\log(1/p)+n)$, where n is the number of vertices in the graph and where p is the probability of an edge. This running time is faster than $O(n\log n)$, the expected running time of the best randomized algorithm of Angluin and Valiant.

12. "Minimizing the Number of Tardy Job Limits under Release Time Contraints" (with R. Shamir), *Discrete Applied Mathematics*, Vol. 28 (1990), 45-57.

Here two single-machine scheduling problems are studied: Minimizing the weighted and unweighted number of tardy units, when release times are present. Fast strongly polynomial algorithms are given for both problems: For problems with n jobs, we give algorithms which require $O(n\log n)$ and $O(n^2\alpha(n))$ steps, for the unweighted and weighted problems respectively. Our results also imply an extension of the family of very efficiently solvable transportation problems, as well as these which are greedily solvable using the ''Monge sequence'' idea.

13. "$O(n \log^2 n)$ Algorithm for the Maximum Weighted Tardiness Problem" (with R. Shamir), *Information Processing Letters*, 31 (1989), 215-219.

We give an $O(n\log^2 n)$ algorithm which solves the scheduling problem of minimizing the maximum weighted tardiness. This improves the classical $O(n^2)$ algorithm of Lawler for that problem. Our algorithm makes a novel use of a data structure which originated in computational geometry.

14. "An Algorithm for the Detection and Construction of Monge Sequences" (with A. Alon, S. Cosares and R. Shamir), *Linear Algebra and its Applications*, 114/115 (1989), 669-680.

This paper describes an efficient algorithm which determines whether a condition due to Hoffman (1963) is satisfied by the cost matrix of a Transportation problem. In case the condition is satisfied, our algorithm generates a permutation of the matrix entries (called a Monge sequence), which allows for the solution of the problem in linear time, by way of a "greedy" algorithm. Previously, the existence and the generation of such sequences had been approached separately for every family of transportation problems. The running time of our algorithm is better than that of the best known algorithm for the transportation problem, and thus it can be used as a preliminary step towards solving such problems without an increase in the overall complexity.

WORKING PAPERS

15. "Simple and Fast Algorithms for Linear and Integer Programs with Two Variables for Inequality" (with J. Naor), Manuscript, April 1991, submitted to *Journal of the Association for the Computing Machinery.*

We present an $O(mn^2\log m)$ algorithm for solving feasibility in linear programs with up to two variables per inequality which is derived directly from the Fourier-Motzkin elimination method. (The number of variables and inequalities is denoted by n and m respectively). The running time of the algorithm dominates that of the best known algorithm for the problem, and is far simpler. We then consider integer programming on monotone inequalities, i.e., inequalities where the coefficients are of opposite sign. This problem includes as a special case the simultaneous approximation of a rational vector with specified accuracy, which is known to be NP-complete. However, we show that both a feasible solution and an optimal solution with respect to an arbitrary objective function can be computed in pseudo-polynomial time, and hence, the problem is weakly NP-complete.

16. "On a Polynomial Class of Nonlinear Integer Optimization Problems," manuscript.

We study the problem of optimizing a nonlinear polynomial in nonnegative integer or continuous variables over box constraints. A class of such problems, the ones with nonnegative coefficients of nonlinear terms, is identified as solvable in (strongly) polynomial time. Another class, that strictly generalizes the first one, is the class of polynomials that have the *bipartition property*. The optimization of quadratic multivariate (up to two variables per term) polynomials with the bipartition property over box constraints is also solvable in (strongly) polynomial time. An algorithm for verifying the bipartition property runs in (strongly) polynomial time. For quadratic polynomials a more efficient algorithm verifies the bipartition property in linear time.

17. "The Empirical Performance of a Polynomial Algorithm for Constrained Nonlinear Optimization" (with S. Seshadri), manuscript, March 1991, submitted to *Annals of Operations Research.*

In [HS90] (paper 8), a polynomial algorithm based on successive piecewise linear approximation was described. The algorithm is polynomial for constrained nonlinear (convex or concave) optimization, when the constraint matrix has a polynomial size subdeterminant. We propose here a practical adaptation of that algorithm with the idea of successive piecewise linear approximation of the objective on refined grids, and the testing of the gap between lower and upper bounds. The implementation uses the primal affine interior point method at each approximation step. We develop special

features to speed up each step and to evaluate the gap. Empirical study of problems of size up to 198 variables and 99 constraints indicates that the procedure is very efficient and all problems tested were terminated after 171 interior point iterations. The procedure used in the implementation is proved to converge when the objective is strongly convex.

18. "Lower and Upper Bounds for the Allocation Problem and Other Nonlinear Optimization Problems," manuscript, August 1989, revised 1991, submitted to *Mathematics of Operations Research*.

We demonstrate the impossibility of strongly polynomial algorithms for the allocation problem, in the comparison model and in the algebraic tree computation model, that follow from lower bound results. Consequently, there are no strongly polynomial algorithms for nonlinear (concave) separable optimization over a totally unimodular constraint matrix. This is in contrast to the case when the objective is linear.

We present scaling-based algorithms that use a greedy algorithm as a subroutine. The algorithms are polynomial for the allocation problem and its extensions and are also optimal for the simple allocation problem and the generalized upper bounds allocation problem, in that the complexity meets the lower bound derived from the comparison model. For other extensions of the allocation problem the scaling-based algorithms presented here are the fastest known.

The algorithms are also applicable to solving with ε accuracy the allocation problem and its extensions in continuous variables and in polynomial time.

19. "A Polynomial Algorithm for the K-Cut Problem" (with O. Goldschmidt), submitted to *Mathematics of Operations Research*.

The k-cut problem is to find a partition of an edge weighted graph into k nonempty components, such that the total edge weight between components is minimum. The status of this problem was open for a number of years, and several closely related problems were proved intractable. Here we prove that this problem is NP-complete for an arbitrary k and its version involving fixing a vertex in each component is NP-hard even for $k=3$. We present a polynomial algorithm for k fixed, that runs in $O(n^{k^2/2-3k/2+4}T(n,m))$ steps, where $T(n,m)$ is the running time of the minimum (s,t)-cut on a graph with n vertices and m edges. This algorithm is of potential significance in reliability, clustering and graph partitioning problems. Although only an extended abstract appeared in print, it has already prompted much related work.

INDEX OF PUBLICATIONS

"A Strongly Polynomial Algorithm for the Quadratic Transportation Problem with Fixed Number of Suppliers" (with S. Cosares), manuscript, August 1990, to appear in Mathematics of Operations Research.

"An Exact Sublinear Algorithm for the Max-Flow, Vertex Disjoint Paths and Communication Problems on Random Graphs," August 1988, to appear in Operations Research.

"Why Should Biconnected Components Be Identified First," to appear in Discrete Applied Mathematics.

"The Use of Heuristics, Cutting Planes and Subgradient Optimization for the Multicovering Problem" (with N. Hall), to appear in European Journal on Operational Research.

"A Polynomial Algorithm for an Integer Quadratic Nonseparable Transportation Problem" (with J. G. Shanthikumar and R. Shamir), to appear in Mathematical Programming.

"On the Impossibility of Strongly Polynomial Algorithms for the Allocation Problem and its Extensions," Proceedings of the Integer Programming and Combinatorial Optimization Conference, May 1990, 261-274.

"Strongly Polynomial Algorithms for the High Multiplicity Scheduling Problems" (with R. Shamir), Operations Research, Vol. 39, No. 4, 1991, 648-653.

"Complex Separable Optimization is Not Much Harder than Linear Optimization" (with J.G. Shanthikumar), Journal of ACM, Vol. 37, No. 4 (1990), 843-862.

"The Complexity of Nonlinear Optimization" (with J. G. Shanthikumar), Lecture Notes in Computer Science, 372, Springer-Verlag, July 1989, pp. 461-472.

"Asymptotically Optimal Linear Algorithms for the Minimum K-Cut in a Random Graph" (with O. Goldschmidt), SIAM J. on Discrete Mathematics, February 1990, Vol. 3, No. 1, pp. 58-73.

"A Fast Perfect Matching Algorithm in Random Graphs" (with O. Goldschmidt), SIAM J. on Discrete Mathematics, February 1990, Vol. 3, No. 1, pp. 48-57.

"Minimizing the Number of Tardy Job Limits under Release Time Contraints" (with R. Shamir), Discrete Applied Mathematics, Vol. 28 (1990), 45-57.

"$O(n\log^2 n)$ Algorithm for the Maximum Weighted Tardiness Problem" (with R. Shamir), Information Processing Letters, 31 (1989), 215-219.

"An Algorithm for the Detection and Construction of Monge Sequences" (with A. Alon, S. Cosares and R. Shamir), Linear Algebra and its Applications, 114/115 (1989), 669-680.

INDEX OF WORKING PAPERS

"Simple and Fast Algorithms for Linear and Integer Programs with Two Variables for Inequality" (with J. Naor), Manuscript, April 1991, submitted to Journal of the Association for the Computing Machinery.

"On a Polynomial Class of Nonlinear Integer Optimization Problems," manuscript, August 1989.

"The Empirical Performance of a Polynomial Algorithm for Constrained Nonlinear Optimization" (with S. Seshadri), manuscript, March 1991, submitted to Annals of Operations Research.

"Lower and Upper Bounds for the Allocation Problem and Other Nonlinear Optimization Problems," manuscript, August 1989, revised 1991, submitted to Mathematics of Operations Research.

"A Polynomial Algorithm for the K-Cut Problem" (with 0. Goldschmidt), submitted to Mathematics of Operations Research.